

**IN THE UNITED STATES PATENT AND TRADEMARK OFFICE**

Application : **09/351,086**  
Applicant(s) : **Nevenka Dimotrova**  
Filed : **7/9/1999**  
Confirmation : **9235**  
T.C./Art Unit : **2623**  
Examiner : **Bui, Kieu Oanh T**  
Atty. Docket : **PHA 23,716**

Title: **METHOD AND APPARATUS FOR LINKING A VIDEO SEGMENT TO  
ANOTHER SEGMENT OR INFORMATION SOURCE**

Mail Stop: **APPEAL BRIEF - PATENTS**  
Commissioner for Patents  
Alexandria, VA 22313-1450

**APPEAL UNDER 37 CFR 41.37**

Sir:

This is an appeal from the decision of the Examiner dated 20 February 2007,  
finally rejecting claims 1-2, 4-20, and 26-27 of the subject application.

This paper includes (each beginning on a separate sheet):

- 1. Appeal Brief;**
- 2. Claims Appendix;**
- 3. Evidence Appendix; and**
- 4. Related Proceedings Appendix.**

## **APPEAL BRIEF**

### **I. REAL PARTY IN INTEREST**

The above-identified application is assigned, in its entirety, to **Philips Electronics North America**

### **II. RELATED APPEALS AND INTERFERENCES**

Appellant is not aware of any co-pending appeal or interference that will directly affect, or be directly affected by, or have any bearing on, the Board's decision in the pending appeal.

### **III. STATUS OF CLAIMS**

Claims 3 and 21-25 are canceled.

Claims 1-2, 4-20, and 26-27 are pending in the application.

Claims 1-2, 4-20, and 26-27 stand rejected by the Examiner under 35 U.S.C. 103(a).

These rejected claims are the subject of this appeal.

### **IV. STATUS OF AMENDMENTS**

No amendments were filed subsequent to the final rejection in the Office Action dated 20 February 2007.

### **V. SUMMARY OF CLAIMED SUBJECT MATTER**

The invention comprises a method and system for determining and accessing ancillary information regarding a feature in a video segment being displayed to a user (FIG. 1, page 5, lines 1-10). Of particular note, this determination of an association allows a user to access information regarding a feature in the video, regardless of whether the original input video 12 includes links (hyperlinks) to this other information.

The determined association may be based on a semantic relationship, visual similarity, scene similarity, event similarity, and so on (page 6, line 13 – page 7, line 9; page 8, line 6 - page 9, line 21). When the user selects the feature within the image, the associated source is accessed, and the associated information is displayed or stored for later viewing (page 7, lines 14-17; FIG. 6, page 14, line 12 – page 15, line 10). Alternatively, the determined information from the other source may be displayed automatically, using, for example, a picture-in-picture (PIP) presentation of available material (page 15, lines 11-28).

As claimed in independent claim 1, upon which claims 2-17 depend, the invention comprises a method for processing video, the method comprising (FIGs. 1-6):

- displaying (18) a sequence of video segments (video-B) at a display (18) of a user (page 5, lines 11-13),

- extracting a feature (20-1) from one or more video segments of the sequence (video B) (page 12, line 18 – page 13, line 4),

- determining (41-43, 51-53) an association (48, 58) between the feature (20-1) and at least one additional information source (12; video A) also including that feature (20-2) (page 6, line 18- page 7, line 9; page 13, line 3 – page 14, line 11); and

- defining (44, 54) a link (48, 58) between the feature (20-1) and the at least one additional information source (video A) to facilitate a display of information (36) from the additional information source based at least in part on a selection (64) by the user of the feature (34) while the one or more video segments are displayed (32) to the user (page 3, lines 10-14; page 14, line 12 – page 15, line 3).

As claimed in independent claim 18, the invention comprises an apparatus (FIG. 1) for processing video, the apparatus comprising:

a display (18) that is configured to display a sequence of video segments (page 5, lines 11-13),

a processor (15) that is configured to (FIGs. 4-6):

extract (40) a feature from one or more video segments of the sequence (page 12, line 18 – page 13, line 4);

determine (44) an association between the feature and at least one additional information source also including that feature (page 6, line 18- page 7, line 9; page 13, line 3 – page 14, line 11); and

direct the display of information from the additional information source based at least in part on a selection (62) by a user of the feature in the first video segment while the one or more video segments are displayed on the display (page 3, lines 10-14; page 14, line 12 – page 15, line 3).

As claimed in independent claim 19, the invention comprises an apparatus (FIG. 1) for processing video, the apparatus comprising:

a processor (15) operative to (FIGs. 2-6):

determine (44, 54) an association (48, 58) between a feature (20-1) in one or more video segments and at least one additional information source (video A) that also includes the feature (20-2) (page 6, line 18- page 7, line 9; page 13, line 3 – page 14, line 11); and

display (64) information (36) from the additional information source based at least in part on a selection (62) by a user of the feature while the one or more video segments are displayed to the user (page 3, lines 10-14; page 14, line 12 – page 15, line 3).

As claimed in independent claim 20, the invention comprises an article of manufacture comprising a machine-readable medium (page 19, lines 5-12) containing one or more software programs which when executed (FIGs. 3-6):

- display (18) a sequence of video segments (page 5, lines 11-13),
- extract (40) a feature from one or more video segments of the sequence (page 12, line 18 – page 13, line 4),
- determine (44) an association (48, 58) between the feature and at least one additional information source that also includes the feature (page 6, line 18- page 7, line 9; page 13, line 3 – page 14, line 11); and
- display (18) information from the additional information source based at least in part on a selection (64) by a user of the feature while the one or more video segments are displayed to the user (page 3, lines 10-14; page 14, line 12 – page 15, line 3).

#### **VI. GROUNDS OF REJECTION TO BE REVIEWED ON APPEAL**

Claims 1-2, 4-10, 17-20, and 26-27 stand rejected under 35 U.S.C. 103(a) over Hjelsvold et al. (USP6,546,555, hereinafter Hjelsvold) and Nagasaka et al. (USP 6,400,890, hereinafter Nagasaka).

Claims 11-16 stand rejected under 35 U.S.C. 103(a) over Hjelsvold, Nagasaka, and Jain et al. (USP 6,463,444, hereinafter Jain).

## **VII. ARGUMENT**

### **Claims 1-2, 4-10, 17-20, and 26-27 stand rejected under 35 U.S.C. 103(a) over Hjelsvold and Nagasaka**

MPEP 2142 states:

"To establish a *prima facie* case of obviousness ... the prior art reference (or references when combined) ***must teach or suggest all the claim limitations***... If the examiner does not produce a *prima facie* case, the applicant is under no obligation to submit evidence of nonobviousness."

### **Claims 1-2, 4-10, 17, and 26-27**

Claim 1, upon which claims 2, 4-10, 17, and 26-27 depend, claims a method that includes extracting a feature from a video segment, determining an association between the feature and an additional information source that also includes that feature, and defining a link between the feature and the additional information source to facilitate a display of information from the additional information source based on a selection by the user of the feature while the video segment is displayed to the user.

The combination of Hjelsvold and Nagasaka does not teach or suggest extracting a feature from a video segment and defining a link between the extracted feature and an associated additional information source to facilitate a display of information from the additional information source based on a selection by the user of the feature while the video segment is displayed to the user.

The final Office action acknowledges that Hjelsvold does not teach extracting a feature from a video segment and defining a link between the extracted feature and an associated information source to facilitate a display of information from the additional information source based on a selection of the feature by the user, and asserts that Nagasaka provides this teaching (Office action page 4, last paragraph). The applicant respectfully disagrees with this assertion.

It is clear from the plain language of claim 1 that the extracted feature is a feature that is able to be selected by a user. Nagasaka does not teach or suggest extracting user-selectable features.

Nagasaka teaches a system and method for characterizing video sequences, so that repeating sequences, such as a repeated commercial, can be identified and optionally skipped (Nagasaka, column 9, lines 16-22). The sequence matching can also be applied to replace repeated segments in a recording, such as the opening sequence of each episode of a weekly or daily program recorded on a DVR, with a pointer to a single copy of the segment, to reduce storage requirements (Nagasaka, column 9, lines 5-16).

To efficiently identify repeated video sequences, Nagasaka teaches defining a feature that characterizes each frame of the sequence, then searching for matching sequences of frame-features. The feature of each frame may be, for example, the average color of each frame, a pattern or texture of the frame, a boundary shape, and so on (Nagasaka, column 5, lines 40-47). A sequence of frames may be represented, for example, as blue-blue-blue-green-blue-red<sup>1</sup>; a matching video segment would have this same sequence of frame-features.

Nagasaka does not teach frame-features that are user-selectable. Of particular note, Nagasaka specifically teaches that a single feature value is assigned to each frame, wherein the determined feature value is quantized to a nominal standard value. (Nagasaka, column 5, lines 36-57). Nagasaki's element 106 in FIG. 2 is a "frame feature extractor", and Nagasaki's FIG. 3 clearly indicates a single feature per frame. Further, minor differences between frame features (A, A', A'') are ignored, so that streams of frames are identified by a single feature (A). In the example of Nagasaki's FIG. 2, the sequence of frames is identified as: feature A for frames between  $t_1$  and  $t_{i-1}$ ; feature B for frames  $t_i$  to  $t_{j-1}$ ; feature C for frames  $t_j$  to  $t_{k-1}$ ; and so on (Nagasaka's Feature Table in FIG. 2).

As taught by Nagasaka, each frame or group of frames is characterized by a feature that is represented as a "feature value", such as an average color 'blue', an average texture 'T23', or an average boundary shape 'circular', each likely being

---

<sup>1</sup> In practice, the average color would be represented as a numerical value, such as the average of the RGB values of the pixels of the image, thereby providing the resolution required to generate a fairly unique sequence of frame-feature values for each particular sequence of frames.

stored as a numeric value to facilitate comparisons of sequences of such features. Such features are not, per se, user-selectable, because the features generally represent an average or characteristic value associated with the entire frame, which particular value may not even appear in the frame. That is, for example, there may not be a single pixel in an image with an RGB value that equals the average RGB value of the frame.

Further, even assuming in argument, that Nagasaka's frame feature can be considered to correspond to the claimed user-selectable feature, Nagasaka clearly does not teach determining an association between the feature and an additional information source.

Nagasaka teaches creating a sequence of features and using this sequence to determine associations between sequences of corresponding video frames. In Nagasaka's example of the feature being an average color of the frame, wherein a sequence may be encoded as blue-blue-blue-green-red-green-blue-blue, Nagasaka does not determine an association between the feature "blue" and an additional information source, because without the particular sequence, the fact that a particular frame has an average color of blue is virtually meaningless.

Because Nagasaka does not teach a user-selectable feature, and because Nagasaka does not teach determining an association between a feature and at least one additional information source also including that feature, the applicant respectfully maintains that the Office action has failed to provide a prima facie case to support the rejection of claims 1-2, 4-10, 17, and 26-27 under 35 U.S.C. 103(a) over Hjelsvold and Nagasaka. Accordingly, the applicant respectfully requests a reversal of this rejection by the Board.



Additionally, MPEP 2143 states:

"THE PROPOSED MODIFICATION CANNOT RENDER THE PRIOR ART UNSATISFACTORY FOR ITS INTENDED PURPOSE"

Even assuming, in argument, that the asserted combination of Hjelsvold and Nagasaka includes each feature of claim 1, the applicant respectfully maintains that this combination will not be satisfactory for its intended purpose.

The intended purpose of the asserted combination of Hjelsvold and Nagasaka is "to facilitate a display of information from the additional information source based on a selection by the user of the feature while the video segment is displayed to the user", as claimed in claim 1.

If a user is provided the ability to select Nagasaka's frame-feature, the applicant respectfully maintains that this selection will not facilitate a display of information from an additional information source. Using the example of an average R-G-B color forming a frame-feature, if the user selects 123-73-245 as the feature, the applicant respectfully maintains that displaying information from one or more sources that may also have a frame with an average R-G-B value of 123-73-245 will generally serve no useful purpose.

The applicant further notes that in the context of finding other sources that may contain a selected feature, such as an image of a particular person, rarely will the frames in each source have equal frame-features as defined by Nagasaka. Only if the entire image in each source is the same, including background, foreground, scale, and so on, will Nagasaka's frame-feature be equal in both sources.

Because the combination of Hjelsvold and Nagasaka will be unsatisfactory for the intended purpose of facilitating a display of information from an additional information source based on a selection by the user of the feature, the applicant respectfully maintains that the rejection of claims 1-2, 4-10, 17, and 26-27 under 35 U.S.C. 103(a) over Hjelsvold and Nagasaka is unfounded, per MPEP 2143, and should be reversed by the Board.

Further, in *KSR Int'l. Co. v. Teleflex, Inc.*, the Supreme Court noted that the analysis supporting a rejection under 35 U.S.C. 103(a) should be made explicit, and that it is "important to identify a reason that would have prompted a person of ordinary skill in the relevant field to combine the [prior art] elements" in the manner claimed:

"Often, it will be necessary ... to look to interrelated teachings of multiple patents; the effects of demands known to the design community or present in the marketplace; and the background knowledge possessed by a person having ordinary skill in the art, all in order to determine whether there was an **apparent reason** to combine the known elements in the fashion claimed by the patent at issue. To facilitate review, this analysis **should be made explicit.**" *KSR*, slip op. at 14 (emphasis added).

The applicant respectfully maintains that there is no apparent reason for combining Hjelsvold and Nagasaka in the manner proposed, other than to parody the elements of the applicant's claims.

Hjelsvold teaches a technique for the display of segments of video based on a classification (payment schedule) of a viewer. Nagasaka teaches a technique for finding matching sequences of frames within video streams.

The Office action asserts that one of skill in the art would be motivated to combine Hjelsvold and Nagasaka "in order to provide the user a sequence of video segments, extracting a feature from one or more video segments, defining a link between the feature and the at least one additional information related to the feature as preferred" (Office action, page 5, lines 8-10). The applicant notes, however, that this asserted motivation is found only in the applicant's disclosure, and not in either Hjelsvold or Nagasaka. The Office has done nothing more than paraphrase and repeat the elements of the applicant's disclosure claims. The Office action offers no apparent reason for combining these references outside of the applicant's teachings.

Because there is no apparent reason to combine a system as taught by Hjelsvold that controls the display of video based on a viewer's payment schedule and a system as taught by Nagasaka that finds matching sequences of frames within video streams, the applicant respectfully maintains that the rejection of claims 1-2, 4-

10, 17-20, and 26-27 under 35 U.S.C. 103(a) over Hjelsvold and Nagasaka is unfounded, per MPEP 2143, and should be reversed by the Board.

#### **Claims 18-20**

The Office action relies solely on the rejection of claim 1 to support the rejection of independent claims 18, 19, and 20 (page 7, lines 4-6).

Because the Office action fails to establish a prima facie case to support the rejection of claim 1, and because there is no apparent reason to combine Hjelsvold and Nagasaka, and because the proposed combination will be unsatisfactory for its intended purpose, the applicant respectfully maintain that the asserted basis for rejecting independent claims 18, 19, and 20 is unfounded, and should be reversed by the Board.

#### **Claims 11-16 stand rejected under 35 U.S.C. 103(a) over Hjelsvold, Nagasaka, and Jain**

#### **Claims 11-16**

Claims 11-16 are dependent upon claim 1. In this rejection, the Office action relies upon the combination of Hjelsvold and Nagasaka for teaching each of the elements of claim 1.

As noted above, the combination of Hjelsvold and Nagasaka fails to teach each of the elements of claim 1. Therefore, the applicant respectfully maintains that the rejection of claims 11-16 under 35 U.S.C. 103(a) that relies on the combination of Hjelsvold and Nagasaka for teaching the elements of claim 1 is unfounded, per MPEP 2142.

### CONCLUSIONS

Because the combination of Hjelsvold and Nagasaka fails to teach each of the elements of the applicant's claims, the applicant respectfully requests that the Examiner's rejection of claims 1-2, 4-20, and 26-27 under 35 U.S.C. 103(a) be reversed by the Board, and the claims be allowed to pass to issue.

Because the combination of Hjelsvold and Nagasaka will be unsatisfactory for its intended purpose, the applicant respectfully requests that the Examiner's rejection of claims 1-2, 4-20, and 26-27 under 35 U.S.C. 103(a) be reversed by the Board, and the claims be allowed to pass to issue.

Because there is no apparent reason to combine Hjelsvold and Nagasaka, the applicant respectfully requests that the Examiner's rejection of claims 1-2, 4-20, and 26-27 under 35 U.S.C. 103(a) be reversed by the Board, and the claims be allowed to pass to issue.

Respectfully submitted

/Robert M. McDermott/  
Robert M. McDermott, Esq.  
Registration Number 41,508  
804-493-0707

**Please direct all correspondence to:**  
Yan Glickberg, Esq.  
Philips Intellectual Property and Standards  
P.O. Box 3001  
Briarcliff Manor, NY 10510-8001  
914-333-9618

#### CLAIMS APPENDIX

1. A method for processing video, the method comprising:
  - displaying a sequence of video segments at a display of a user,
  - extracting a feature from one or more video segments of the sequence,
  - determining an association between the feature and at least one additional information source also including that feature; and
  - defining a link between the feature and the at least one additional information source to facilitate a display of information from the additional information source based at least in part on a selection by the user of the feature while the one or more video segments are displayed to the user.
2. The method of claim 1 wherein
  - defining the link includes retrieving the link from a memory based on an identification of the feature.
3. (Canceled)
4. The method of claim 1 wherein
  - the additional information source includes an additional video segment that also includes the feature.
5. The method of claim 4, including
  - switching from display of the first video segment to display of the additional video segment.
6. The method of claim 4, including
  - displaying the additional video segment at least in part in a separate portion of a display which also includes at least a portion of the one or more video segments.

7. The method of claim 1 wherein  
the feature includes a video feature.
8. The method of claim 7 wherein  
the video feature includes at least one of:
  - a frame characterization,
  - a face identification,
  - a scene identification,
  - an event identification, and
  - an object identification.
9. The method of claim 1 wherein  
the feature includes an audio feature.
10. The method of claim 9, including  
combining an audio signal corresponding to the audio feature with an audio  
signal associated with the first video segment.
11. The method of claim 9, including  
converting an audio signal corresponding to the audio feature into a textual  
format that is displayed with the first video segment.

12. The method of claim 9, including  
    separating at least a portion of the video segment into audio categories  
including one or more of:  
    single-voice speech,  
    multiple-voice speech,  
    music,  
    silence, and  
    noise,  
in order to extract the audio feature therefrom.
13. The method of claim 9, wherein  
    the audio feature includes at least one of:  
        a music signature extraction,  
        a speaker identification, and  
        a transcript extraction.
14. The method of claim 1, wherein  
    the feature is a textual feature.
15. The method of claim 14, including  
    displaying information corresponding to the textual information as an overlay  
on a display of the first video segment.
16. The method of claim 1, wherein  
    the feature includes at least one multi-dimensional feature vector extracted  
from a portion of the video segment using a feature extraction technique.
17. The method of claim 1, wherein  
    determining the association includes determining a similarity measure using a  
clustering technique.

18. An apparatus comprising:

a display that is configured to display a sequence of video segments,

a processor that is configured to:

extract a feature from one or more video segments of the sequence;

determine an association between the feature and at least one

additional information source also including that feature; and

direct the display of information from the additional information source

based at least in part on a selection by a user of the feature in the first video segment while the one or more video segments are displayed on the display.

19. An apparatus for processing video, the apparatus comprising:

a processor operative to:

determine an association between a feature in one or more video segments and at least one additional information source that also includes the feature; and

display information from the additional information source based at least in part on a selection by a user of the feature while the one or more video segments are displayed to the user.

20. An article of manufacture comprising a machine-readable medium containing one or more software programs which when executed:

display a sequence of video segments,

extract a feature from one or more video segments of the sequence,

determine an association between the feature and at least one additional information source that also includes the feature; and

display information from the additional information source based at least in part on a selection by a user of the feature while the one or more video segments are displayed to the user.



21-25 (Canceled)

26. The method of claim 1, including

storing the link to facilitate subsequent display of the information from the additional information source.

27. The method of claim 1, including

combining the link and the video segment to create a hyperlinked video segment.

**EVIDENCE APPENDIX**

No evidence has been submitted that is relied upon by the appellant in this appeal.

**RELATED PROCEEDINGS APPENDIX**

Appellant is not aware of any co-pending appeal or interference which will directly affect or be directly affected by or have any bearing on the Board's decision in the pending appeal.